Deconstructing orchestration load: comparing teacher support through mirroring and guiding

Ishari Amarasinghe[1] · Davinia Hernández-Leo[1] · H. Ulrich Hoppe[2,3]

Abstract

Under the notion of "CSCL scripts", different pedagogical models for structuring and supporting collaboration in the classroom have been proposed. We report on a practical experience with scripts based on the Pyramid collaborative learning flow pattern supported by a specific classroom tool and a teacher-facing dashboard that implements mirroring and guiding support. The input data of our analysis stems from recordings of classroom interactions guided by several teachers using the PyramidApp with different levels of teaching support. For the analysis, we introduce a specific coding scheme enabling a quantitative comparison and deeper analysis using epistemic network analysis. The results show that the guiding support enabled teachers to perform more orchestration actions, more targeted interactions and to make more announcements to the class (regarding time, phase transitions, and students' activity participation) when compared to the mirroring support. Teachers' actionable differences observed under the mirroring and guiding support directed us to deconstruct the notion of orchestration load into different facets and to discuss how different support provisions correspond to the different facets of orchestration load.

Keywords CSCL scripts · Orchestration · Dashboards · Learning analytics · Collaboration · Epistemic network analysis

✉ Ishari Amarasinghe
  ishari.amarasinghe@upf.edu

  Davinia Hernández-Leo
  davinia.hernandez-leo@upf.edu

  H. Ulrich Hoppe
  uh@rias-institute.eu

[1] Universitat Pompeu Fabra, Barcelona, Spain

[2] University of Duisburg-Essen, Duisburg, Germany

[3] RIAS Institute, Duisburg, Germany

Springer

Introduction

Teachers have a central and decisive role in the context of classroom collaborative learning. In formal co-located collaborative educational settings, teachers usually engage in designing collaborative learning activities, executing those designs in the classrooms while coordinating, monitoring, evaluating and providing support to students when required (Sharples, 2013). However, the execution of such learning activities may not always unfold according to the original plan as extraneous activities that were not predicted during the activity design time may create deviations. Such unpredicted yet unavoidable incidents that occur during activity enactment demand teachers to make adequate design decisions in real-time and to adapt the learning design on-the-fly to attain fruitful learning outcomes, and to meet students' expectations (Roschelle et al., 2013).

In the context of collaborative learning, the notion of orchestration has been put forward to describe how teachers productively coordinate and manage classroom activities at different scales, e.g., individual, small group, and class-wide activities, under multiple constraints in real-time (Dillenbourg et al., 2009). Teacher centrism is a key feature within the concept of orchestration in which the role of teachers' is not conceived only as the one of a *guide on the side* but rather as a *conductor*, who manages and guides the learning activities in a productive direction (Dillenbourg & Jermann, 2010; Sharples, 2013). Orchestration of collaboration is known to be a demanding task for the teachers as it requires effort to balance epistemic and social aspects of collaboration while taking into account other constraints, e.g., time, space, discipline, all of which emerge within classroom spaces (Cuendet et al., 2013).

Commonly referred to as "orchestration technology", an extra layer of technology can be introduced within technology-enhanced classroom learning spaces to support teachers in orchestrating collaboration (Prieto et al., 2018). From a usability perspective, orchestration technologies require taking into account the "usability at the classroom level" in the sense that the classroom is seen as the user environment in which teachers operate under complex constraints in their management of learning activities (Dillenbourg et al., 2011). But what kind of design features of orchestration technology do teachers find useful in regulating collaboration? How do different support provisions influence teachers' orchestration load? These are important aspects that remain to be explored within the field of technology-enhanced learning research.

The notion of orchestration load is still a "fuzzy" concept (Prieto et al., 2015), and different definitions can be found in the literature. For instance, Cuendet et al. (2013), define orchestration load as the effort necessary for the teacher–and other actors–to conduct learning activities at the class-wide level. In Prieto et al. (2018), orchestration load is described as the effort a teacher spends in coordinating multiple activities and learning processes. In Dillenbourg (2015), orchestration load is described as a factor that includes both the workload (or the energy that teachers need to put in to monitor a learning situation, to perform adaptations, etc.) and the cognitive load (the amount of cognitive resource required to process information, to think, and to take actions, etc.). As described in Prieto et al. (2015) existing studies recognize orchestration load as a concept that contributes to both physical and cognitive effort teachers are required to put in when regulating learning activities in real-time.

Despite the fuzziness associated with the definition of this concept, it is important to recognize and appreciate different types of load that teachers may experience during classroom orchestration as important factors without disregarding them in a negative sense. For

instance, a teacher evaluating multiple student answers during a CSCL activity in order to detect mistakes or to provide immediate feedback or to prepare final debriefing (aspects related to the epistemic aspect of collaboration) (Dillenbourg, 2015; Martinez-Maldonado et al., 2015) can contribute to the *content load* of the teachers. In terms of cognitive load theory (Sweller, 2020), processing of new information in activities of short duration, can contribute to the *intrinsic load* of the teachers as it requires cognitive processing of information in working memory. Moreover, diagnosing certain students' deviations from group activity (related to the social aspect of collaboration) and other constraints, e.g., effective use of available time, can add to the *orchestration load* of the teachers. Further, when additional support in the form of orchestration technology is available, understanding the technological support itself to take relevant pedagogical actions can add to the *cognitive load* of the teachers (Sharples, 2013). All types of load, i.e., *content, orchestration* and *cognitive load*, reflect the valuable and attentive cognitive processing teachers are required to engage in real-time when orchestrating CSCL activities, which are essential to achieving the intended learning goals.

However, most of the existing studies either disregard the aforementioned aspects or refer to the notion of orchestration load as a black box without exploring it in detail due to several reasons. For instance, difficulties related to understanding how orchestration load emerges, factors that influence it in authentic educational situations, difficulties associated with grounding this notion in empirical evidence, and the unavailability of standard measurements to quantify orchestration load are to name a few. However, from a design perspective, design processes that disregard orchestration load for tools meant to support teachers produce technologies that may introduce an additional burden instead of supporting and simplifying activity regulation (Prieto et al., 2018; Sharples, 2013).

To this end, the goal of the study is to deconstruct the notion of the orchestration load and to understand its multifaceted elements, which will facilitate broadening our understanding of this complex notion. To achieve this goal, we have modelled teachers' orchestration actions under different supporting conditions in authentic CSCL situations, and then used the modelled actionable differences to derive different facets of the notion of orchestration load.

CSCL Scripts

In CSCL, group learning can be structured pre-emptively using collaboration scripts (Dillenbourg, 2002). By proposing an activity sequence and allocating roles to students with specific duties and responsibilities, scripts aim to trigger certain types of beneficial collaborative learning interactions between students (Kobbe et al., 2007). Several studies have reported the effectiveness of using scripts to achieve productive learning outcomes in collaboration (Radkowitsch et al., 2020; Rummel & Spada, 2005).

Although scripts provide a structure for collaboration that favors learning, disturbances that may occur during its enactment can cause deviations from the original plan (Dillenbourg & Tchounikine, 2007). For instance, consider the deployment of the pattern-based Pyramid script (Hernández-Leo et al., 2019) in a classroom context. This script structures the collaborative learning flow in such a way as to encourage students to reach a consensus within a number of phases that occur one after the other, following a pyramid structure. The pattern integrates activities from multiple social planes, i.e., individual, group, and class-wide levels as described below. First, learners start to solve a given problem individually. Then they join in small groups (usually pairs) to share their solutions and to agree

on common solutions. Later small groups are merged, forming increasingly larger groups as the activity flow advances. The increasingly larger groups make up the Pyramid structure. Thus the Pyramid script mediates learning and reflection within different stages of the script. It aims to provide opportunities for all learners to express their solutions and to discuss their ideas with peers.

However, in order to attain fruitful collaboration and therefore learning, this pattern expects to contribute and sustain participation from the beginning until the end of the consensus-building process. Lack of individual motivation and participation in different phases of the Pyramid script can reduce students' ability to reach a consensus at the end. This will result in a less productive collaborative learning experience for motivated students. Under a lack of expert monitoring, there are chances that the script may lead students to reach a potentially misleading consensus that is not aligned with the pedagogical intentions of the teachers. Moreover, as groups work in parallel, they advance the activity flow to different degrees. This may create periods of idle time for faster groups, which in turn might lead to off-task behavior in the classroom, whereas slower groups may require more time to produce collaboration outcomes. On the one hand, such eventualities can impede achieving beneficial learning outcomes and require a teacher's immediate intervention for further guidance, script adaptation, and regulation (Rodríguez-Triana et al., 2015). On the other hand, it is difficult and oftentimes not feasible for teachers to constantly distribute their attention across different social planes to track progress and to elucidate the required interventions (van Leeuwen, 2015).

Orchestration as a task for teachers realizing CSCL activities

As previously mentioned, the orchestration metaphor captures the complex set of coordination actions teachers are required to handle on different social planes in real-time in highly constrained learning situations (Roschelle et al., 2013). For instance, orchestration actions that occur at the teacher-individual student level may include answering individual questions that request task-related clarifications. Orchestration actions that occur at the teacher-classroom level may include praising and criticizing activities (for positive and negative behavior of students), surveying and perception activities (to diagnose collaboration), giving directions, debriefing activities (to summarize activity outcomes) as well as announcements to the whole class (related to remaining time, phases of the activity and participation). In addition to the above, in scenarios where additional support from orchestration technology is available, teachers may also engage in understanding the technological support available and make decisions on effective ways of using technologies to monitor and diagnose collaboration.

As described in Soller et al. (2005), managing collaboration in real-time can be described as a cyclic activity, in which the current state of the interactions is continuously compared against the desired state in order to detect discrepancies. Detection of deviations will call for remedial actions by the teachers to achieve the goals and objectives of the learning situations. Assessing learning situations in real-time in order to detect deviations and taking relevant actions is a demanding task that adds to the orchestration load of the teachers. As described previously, not only actions that are occurring at individual, group and class levels but also the assistive orchestration technologies can contribute to the orchestration load, as teachers are required to employ their cognitive resources to understand the technological support provided and to make the best use of those technologies.

Teacher dashboards

Learning Analytics (LA) dashboards can be conceived as "single displays that aggregate different indicators about learner(s), learning process(es) and/or learning context(s) into one or multiple visualizations" (Schwendimann et al., 2016; Verbert et al., 2014). Recently a growing research interest towards provisioning teacher-facing LA dashboards to support teachers has been observed (Martinez-Maldonado, 2019; van Leeuwen & Rummel, 2020; Wise & Jung, 2019). These dashboards visualize pertinent learner-educational platform interaction data and aim to support teachers in monitoring and taking informed pedagogical actions (Amarasinghe et al., 2020). In other words, LA dashboards can be used to support the regulation loop of orchestration (Soller et al., 2005). By aligning LA with the pedagogical intentions documented in the learning design, dashboards can be used to promote critical moments or activity deviations. Using checkpoint and process analytics teachers can look for specific patterns in the data at predefined time points, e.g., successful and unsuccessful engagement patterns, in order to provide relevant feedback for students to enhance their interactions (Lockyer et al., 2013).

Such supporting tools can be assigned into different categories based on the granularity of the support available. As described in Soller et al. (2005) *mirroring tools* visualize learners' interactions when engaged in online learning systems. The end-users of the mirroring tools are expected to diagnose the learning situation, e.g., collaboration, based on the given information, and to decide remedial actions. In contrast, *guiding tools* visualize relevant information but also recommend and guide end-users to take remedial actions to enhance the learning situation. A recent review conducted in van Leeuwen and Rummel (2019) a similar categorization of orchestration tools, i.e., mirroring, alerting, and advising tools, have been proposed. van Leeuwen and Rummel (2019) described the mirroring tools as systems that provide information but do not facilitate the interpretation of information. Alerting tools facilitate the interpretation of information by alerting the teachers about critical events that occur during collaboration. Advising tools recommend teachers to take certain remedial actions. The authors have shown that the teacher-facing dashboards provide advising support and help teachers to detect problematic groups often in a simulated learning environment when compared to the dashboards that provide mirroring support.

However, there is a dearth of studies that compare how these different types of support influence teachers' orchestration actions and orchestration load in authentic collaborative learning scenarios (Wise & Jung, 2019; Martinez-Maldonado, 2019; van Leeuwen and Rummel, 2019). Moreover, despite the increased amount of research attempts to deploy teacher-facing dashboards to support teachers, recent studies have also shown that the adoption of LA dashboards and other LA tools in general within teaching practice is still low (Prieto et al., 2019; Schwendimann et al., 2016). Some studies have also raised questions regarding the deficiencies associated with the design process of such technologies, e.g., lack of inter-stakeholder communication (practitioners, students), and their involvement during LA tool design processes (Prieto et al., 2019). Shum et al. (2019a) pointed out that the design of LA tools should go beyond the technological and pedagogical principles and require incorporating human factors questioning why and how such tools will be used in everyday practices. When considering such human-centered design perspectives within the context of CSCL, the notion of orchestration load plays an important role, because tools and technologies that add to this load are

not likely to integrate into everyday teaching practice as they do not help to augment teachers' actionability. Hence modelling teachers' orchestration actions under different supporting conditions can help to elucidate how different support provisions correspond to an increased or decreased orchestration load and to derive different facets of orchestration load. This can in turn shed light on the types of orchestration technologies that lower the orchestration load of the teachers and the different components that the black box of orchestration load entails.

Research question and expectations

We studied how different support provisions facilitate teachers' orchestration actions and influence their orchestration load in authentic scripted classroom collaborative learning contexts. Pyramid pattern based CSCL activities were deployed in classrooms using the PyramidApp tool, and teachers were given access to a LA dashboard to facilitate orchestration. The dashboard implemented two different types of support, namely mirroring and guiding. In the mirroring support, the interpretation of information and the dashboard use was decided by the teacher without additional guidance. In the guiding support teachers had access to the same dashboard, but they were directed to take actions via an alerting mechanism that flagged critical moments in collaboration. As a control condition, we also included a no dashboard condition. As the name implies in this condition teachers did not have access to a teacher-facing dashboard. The interpretation of collaboration was based on classroom cues, e.g., teacher's observations and questions raised by students.

Teachers' orchestration actions in the three conditions were modelled using Epistemic Network Analysis (ENA) (Shaffer et al., 2016). Using a mixed-method approach we then triangulated the results of the ENA with teachers' subjective perceptions of the different supporting options. Teachers' subjective estimates of the cognitive load experienced under different support provisions were collected using a questionnaire. This measure did not capture the differentiation between load types but rather reflected the overall effort teachers were required to make during orchestration. The central research question addressed in this study is: *how do mirroring and guiding support provided in teacher-facing LA dashboards influence orchestration load?*

Our expectations regarding different supporting provisions were the following: in the no dashboard condition, we expected that teachers may have less awareness and less control over the CSCL activity. Due to lack of access to relevant information, we expected that in this condition teachers may face difficulties in focusing on both epistemic and social aspects of the learning situation and would not be able to make announcements to the class regarding participation levels, phase transitions and remaining time. We expected that due to a reduction in cognitive activities teachers were required to engage in, they would experience the lowest cognitive load under the no dashboard condition when compared to the other two conditions, i.e., mirroring and guiding support.

We expected that under the mirroring and guiding support conditions teachers would have a high awareness and control over the activity when compared to the no dashboard condition. Due to their access to relevant information we expected that in both mirroring and guiding conditions teachers might not face difficulties in focusing on both epistemic and social aspects of the learning situation and would be able to make announcements to the class regarding the levels of students' activity participation, phase transitions and remaining time.

However, regarding the mirroring support we expected that teachers would perform a smaller number of dashboard interventions when compared to the guiding support condition. The reason is that making sense of the information presented in the dashboard for evaluating the learning situation, formulating goals, understanding the support, and deciding on relevant interventions is left to the teacher, which is demanding in real-time. We assumed that this situational demand would result in relatively a high cognitive load for the teachers when compared to the guiding support.

In contrast, under the guiding support, we expected that teachers would perform a higher number of dashboard interventions since automatic alerts were used to signal critical events, and that required teacher interventions. We expected that the alerts would provide additional support for evaluating the learning situation, formulating goals, and taking action at individual, group and class levels. Moreover, as alerts were used to guide teachers' actions, we expected that they might devote fewer cognitive resources to understand the support but might devote more cognitive resources to evaluate the epistemic and social aspects of the learning situation (high focus). Due to the aforementioned reasons, we assumed that teachers would experience a relatively low cognitive load under guiding support when compared to the same load experienced under mirroring support. Having formulated the aforementioned expectations about the three conditions, we conducted a case study with six-teachers following a within-subjects design.

The following sections of the paper are organized as follows: First, we present details about our authentic CSCL study. Next we provide the study results followed by a discussion about the results. Finally, the limitations of our study, conclusions derived and directions for future research are outlined.

Methods

Technical Facilitation (PyramidApp and Teacher-facing Dashboard)

In this study a web-based tool called PyramidApp that implements a special version of the pyramid script was used to deploy collaborative learning activities in the classrooms (Manathunga and Hernández-Leo, 2018). The tool provides an activity authoring space and a teacher-facing LA dashboard for teachers as well as an activity enactment space for students.

Figure 1 shows the PyramidApp's authoring user-interface. When authoring a Pyramid activity, teachers are required to enter the question to be answered by the students and configure the following parameters according to the unique requirements of their classrooms: (1) size of the class; (2) size of small groups; (3) number of levels in the pyramid; (4) number of participants per pyramid; (5) duration for answer submission and subsequent group levels; (6) keywords teachers expect to see in students' answers (up to 10 maximum). Apart from the aforementioned parameters, teachers also have the option to configure automatic alerts that can inform critical moments related to collaboration (see Table 1).

In PyramidApp, collaboration among students is facilitated across different Pyramid levels as follows. After having logged into PyramidApp, students are required to submit an answer individually to the given problem. After submitting their answers, students need to wait until the predefined time for answer submission expires. At the end of the individual answer submission phase, students are automatically randomly allocated into small groups. Within small groups, students can see the answers submitted by their group members along

Fig. 1 Authoring user interface of the PyramidApp, basic parameter configuration (top-left), time related configurations (bottom-left), alert configuration (top-right), keyword configuration (bottom-right)

with a voting mechanism, which offers the opportunity for them to vote on each answer. An integrated chat facilitates discussion among students within their respective groups. Small groups are later merged into larger groups in which highly voted answers at the small group levels are shown to students for further evaluation. Figure 2 shows a screenshot of the PyramidApp as students use it during a group phase. The tool enables potentially productive learning with features meant to elicit desired fruitful student actions. This design is aligned with related technical approaches for supporting physically co-present learning activities, such as classroom response systems (Schell et al., 2013) or backchannel systems for collaborative classrooms (Gehlen-Baum et al., 2014).

As explained previously the teacher-facing LA dashboard is meant to support teachers in orchestrating PyramidApp based collaboration. The dashboard implements two different types of support: mirroring and guiding based on the additional guidance provided using alert mechanisms. It consists of two tabs, namely, *Responses Related* and *Participation Related*. As shown in Fig. 3, the *Responses Related* tab displays the individual answers submitted by students and highly voted answers based on activity at the small group level and the final selected answers based on activity at the large group level. Keywords detected in students' answers (using a custom keyword searching algorithm) were highlighted (shown in green).

The *Participation Related* tab (see Fig. 4) displays the participation of groups. Within this tab, the level of activity in groups was visualized using boxes. For a given group, a larger box shows the voting participation percentage of the members of the group and

Table 1 Dashboard alerts

Name of the alert	Criteria or the rule	Intended use
No keywords	When a keyword is not detected within the first 20% of students' answers	This aims to ensure teachers know whether the answers submitted by the students align with the teacher's expectations
Answer submission skipped	When 20% of the logged in students skipped answer submission	This aims to provide a hint to the teacher if students start skipping answer submission
More time for submission	When 50% of the logged in students have not submitted answers and 50% of the submission time has passed	This aims to provide a hint to the teacher if the majority of the students require more time to submit answers
More time for voting	When 50% of the Pyramid groups have not finished voting and 50% of the voting time has passed	This aims to provide a hint to the teacher if the majority of the students require more time to vote on answers

Fig. 2 User interface of the PyramidApp, voting space (left), discussion space (right)
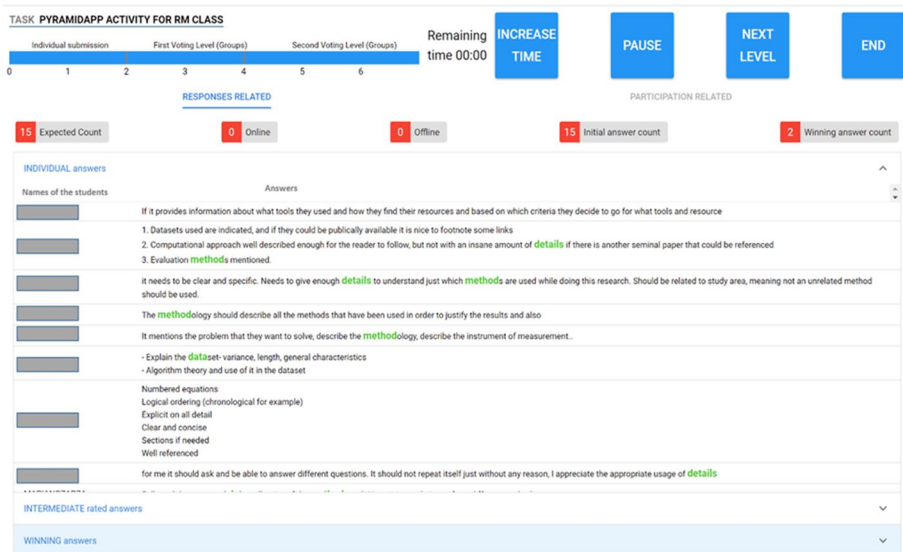
Fig. 3 Information presented in the *Response Related* tab of the dashboard

a small box shows the number of messages posted in the chat, indicating whether they engaged in discussion. The voting percentage and number of messages posted in groups are updated in real-time. Based on voting and discussion participation, groups were classified under two categories: satisfactory and unsatisfactory participation. This group classification aimed to provide a glimpse into students' participation in the activity at a given moment. Upon touching group boxes in the interface, teachers can obtain more details about groups, e.g., names of the group members, answers to be voted in a given group,

Fig. 4 Information presented in the *Participation Related* tab of the dashboard. Two groups are shown under the unsatisfactory participation category due to their lack of participation in the discussion (indicated as a zero)
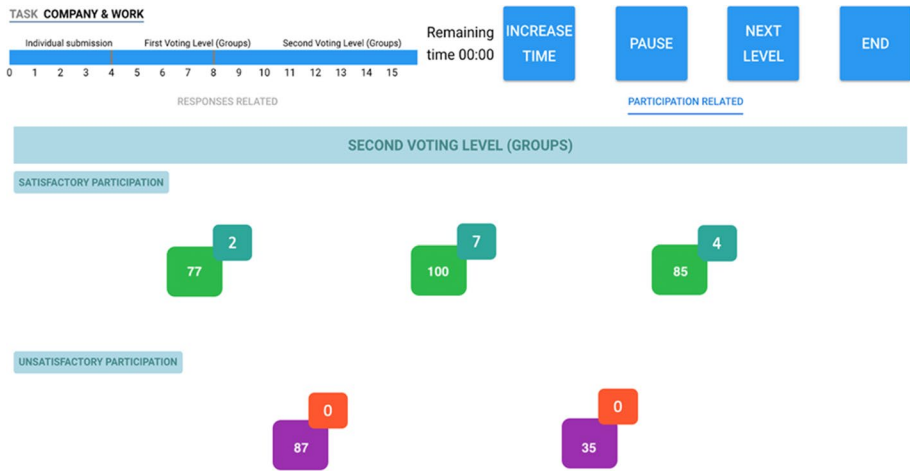
students' participation in the chat. Teachers were also able to intervene in students' chat by posting predefined messages to groups in real-time.

A timeline visualization and a remaining time countdown were added to the dashboard to make teachers aware of the real-time progression of the activity and the remaining time (see Fig. 4 top-left). Four controls were added to the dashboard as buttons to allow teachers to modify the script manually during the runtime of the activity (see Fig. 4 top-right). For instance, the increase time button allows teachers to increase time for the currently active pyramid level, whereas the pause button allows pausing and resuming the activity at any moment, and the next level button allows for skipping intermediate group levels in the pyramid script. Finally, the end button allows for stopping the progression and exit the activity whenever required. The teacher's dashboard actions (that were taken by using control buttons or as a response to a dashboard alert) were communicated to the students as a notice appearing on top of the PyramidApp user-interface. Teacher dashboard actions and students' enactment activity data are logged in the PyramidApp database.

Study participants and experimental design

Following a within-subject design, six higher-education teachers (three females) from a public university in Spain participated in our study. All six teachers had prior experience with the PyramidApp. However, none of them had experience with using dashboard applications to orchestrate collaboration. Before the experiments, each teacher was introduced to the features and functionalities of the dashboard and trained on how to use it. Each teacher participated in the training for around 45 min to one hour.

Each teacher conducted three different collaborative learning sessions addressing the three conditions that we were interested in (see Table 2). The design of each collaborative learning activity varied based on the teacher's requirements to conduct CSCL activities in their classrooms and the time available (see Table 2). As shown in Table 2, teachers A, B, and C followed the following order: no dashboard condition, mirroring

Table 2 A summary of collaborative learning activities conducted, reflecting the order of activities

Teacher ID	Order of sessions and details			Question given to students in all three sessions
	No dashboard condition	Mirroring condition	Guiding condition	
A	Time: 9 min Students: 30	Time: 9 min Students: 71	Time: 9 min Students: 48	Share knowledge about Information Technology (IT) companies and jobs
B	Time: 6 min Students: 36	Time: 6 min Students: 31	Time: 6 min Students: 28	Share knowledge about research writing
C	Time: 6 min Students: 17	Time: 6 min Students: 16	Time: 6 min Students: 19	Share knowledge about presentation skills
Teacher ID	Order of sessions and details			Question given to students in all three sessions
	No dashboard condition	Guiding condition	Mirroring condition	
D	Time: 13 min Students: 31	Time: 13 min Students: 24	Time:13 min Students: 19	Share knowledge about ethics in IT
E	Time: 9 min Students: 7	Time: 9 min Students: 21	Time: 9 min Students: 51	Share knowledge about video production
F	Time: 15 min Students: 13	Time: 20 min Students: 19	Time: 20 min Students: 12	Share knowledge about distributed systems and ethics

condition, and guiding condition whereas teachers D, E, and F followed the following order: no dashboard condition, guiding condition, and mirroring condition. The total time allocated for each activity, the number of students who participated, and the questions proposed by the teachers for different activities are presented in Table 2.

Data Collection

All experiments were video recorded. Apart from logging the teachers' dashboard actions, we also collected screen-captured data (audio and video) from the dashboard tablet. An author of this study transcribed the video recordings to create a dataset that included timestamped information regarding teachers' actions. Transcribed video data and screen-captured data were then merged along the timestamps to create a single dataset that described each teacher's actions during each collaborative learning session. At the end of each session, teachers were given two post-activity questionnaires. One questionnaire focused on teachers' perceived experience regarding the CSCL activity and support provided. In order to collect teachers' perceived cognitive load, we followed the guidelines from previous research (Prieto et al., 2015) and provided them another questionnaire. They were asked to rate their perceived cognitive load on a scale from 1 to 20 (1 low and 20 high). Figure 5 shows the technical setup used for experimentation.

Fig. 5 A teacher using the dashboard and students' enactment in PyramidApp (top), data collection in a classroom session (bottom)

Coding teachers' actions

To analyze the behavioral data collected we defined a coding scheme (following iterative refinements) to code the teacher's actions. The codes that we propose are in alignment with the notion of orchestration and the particular CSCL script being orchestrated. At first, we came up with a detailed coding scheme that consisted of nineteen codes to code teachers' actions. However, we realized that some of those codes, e.g., reflection, are not directly observable in the video recordings and are more related to cognitive aspects. As we did not collect data to interpret such cognitive aspects, we improved the initial coding scheme eliminating such codes and including only the codes that reflected teachers' observable behaviors. The number of codes also influences the visual interpretability of the ENA networks, which span over the codes as nodes. It is important to our characterization that the unit of analysis captures the interconnections between codes. The codes are predefined constructs that shape the possible inter-relations. A smaller number of less sparse codes is thus required for better comparability and (visual) interpretability. This was another more technical reason for us to define agregated codes that captured a number of teachers' actions (see Table 3). Accordingly, we simplified our coding scheme to contain only seven codes in total that captured teachers' observable behavior when orchestrating collaboration.

The first four codes shown in Table 3 were used to code the data obtained from all three experimental conditions (e.g., no dashboard, mirroring and guiding support). The last three codes shown in Table 3 were only applied to code teachers' behavior during mirroring and guiding support conditions, in which the teachers used the dashboard to orchestrate collaboration. An author of this study and an external researcher coded the dataset using 1's and 0's for each binary code, thus indicating the presence and absence of the codes. There was high agreement between the coders (Cohen's Kappa = 0.95, $p < 0.005$), and the relatively few disagreement were resolved through discussion. We then applied ENA techniques to model the structure of connections between coded elements in the data.

Modelling teacher's actions using ENA

ENA is an analysis and modelling approach that combines coding techniques used in qualitative studies with statistical modelling (Shaffer et al., 2016; Shum et al., 2019b). ENA quantifies the connections between pre-specified codes in discourse based on co-occurrence within a sliding window and visualizes the structure of connections as networks with weighted edges over time (Shaffer et al., 2016). The relations (edges) are based on co-occurrences, and the edge weights on the multiplicity of such co-occurrences. The edge weight is visualized as the thickness of the connection. Since co-occurrences are counted within a window of proximity and this window slides over a given sequence of codes in temporal order, the interlinking is time-dependent thickness, which is different from "bag of words" approaches in linguistic analyses computed over whole documents, which would offer a cross-sectional view instead. In this sense, ENA is able to account for temporality in interaction and discourse data. This responds positively to the injunction for including temporal aspects in the analysis of learning processes (Knight et al., 2017; Reimann, 2009). This time-sensitive characteristic is an advantage over aggregated frequency-based measures that do not capture potentially important sequential and temporal co-occurrences associated with learning processes (Saint et al., 2020). Csanadi et al. (2018) have shown that such cross-sectional coding-and-counting strategies can produce misleading insights.

Table 3 Codes defined to describe teachers' actions

Condition	Code	Definition
Applicable to all three conditions (no dashboard, mirroring and guiding)	Teacher individual interaction	Teacher responds/answer to specific questions raised by individual students (unidirectional)
	Teacher class interaction	This code captures the bidirectional interactions between teachers and the whole class. Examples: Teacher requests information from the class (Surveying) Teacher provides directions to the class on how to use the PyramidApp or describes the task (Giving directions) Teacher discusses the final selection of answers at the end of the class (debrief) Teacher advises lower participation students (criticize)
	Announcements to class	Teacher makes announcements to the class regarding: Time remaining for the activity Phase transitions of the Pyramid script Levels of students' participation in the activity
	Teacher perception	This includes the following two behaviors: Teacher is looking at individual student devices (e.g., mobile or desktop monitors) Teacher is looking at the task projection
Applicable only to mirroring and guiding conditions	Check responses tab	This code summarizes the following actions by the teacher within the *Responses Related* tab of the Dashboard: Scrolling answers received from individual students Scrolling highly rated answers at the group level Checking other information (e.g. online & offline counts, number of answers, etc.)
	Check participation tab	This code summarizes the following actions by the teacher within the *Participation Related* tab of the Dashboard: Checking information related to satisfactory and unsatisfactory voting participation of groups Opening a group box and scrolling the chat messages posted by the students Opening a group box and checking the names of the group members

Table 3 (continued)

Condition	Code	Definition
	Dashboard interventions	This code summarizes the following dashboard interventions by the teacher:
		Opening a particular group box and posting messages to groups
		Rating on behalf of student groups
		Use of *Next Level* button to move to the next level of the activity
		Use of *Increase time* button to increase the time for the activity
		Use of *End* button to end the collaborative learning activity before reaching the original designed time
		Use of *Pause* button to pause the script

In the context of CSCL, ENA has been applied for a variety of modelling purposes, ranging from models of students' actions in collaborative learning settings (Oshima et al., 2019; Sung et al., 2019), contributions within collaboration discussion spaces (Ma et al., 2019), to generating visualizations to support teachers' interventions (Herder et al., 2018) as well as feedback in co-located collaborative situations (Shum et al., 2019b).

In our study, ENA has been used to model teachers' actions captured in co-located CSCL situations. During the modelling process, teachers who conducted collaborative sessions across the three different experimental conditions (e.g., no dashboard, mirroring and guiding) were targeted as the unit of analysis. This means that we use the network representations generated through ENA to characterize the behavior of a teacher in a given classroom session. The experimental conditions were set as the conversation variable. Our analysis is based on a sliding window size ("stanza") of three. The size of the moving stanza window was chosen after a qualitative assessment of the dataset, with the goal of capturing meaningful connections in discourse (Siebert-Evenstone et al., 2017). A window size of three also excludes longer distance dependencies but goes beyond bigram Markov models, which would correspond to a window size of two. Basically, the moving stanza window method moves over data and counts the connections between codes that occur within the size of the given window. This process of accumulating connections is repeated for each unit of analysis (teacher-session) resulting in a matrix of adjacency vectors that represent units in rows and connections in columns. ENA performs a dimensionality reduction of the data using singular value decomposition (SVD) to determine a reduced set of new dimensions that preserves maximum variance among the units. ENA also calculates a centroid for a given network model, which is the arithmetic mean of the edge weights. Hence this centroid summarizes the network as a single point and provides a summarized visualization for each unit's network in the projection space.

We believe that ENA is appropriate for our modelling task for several reasons: First ENA takes into account the temporality of teachers' actions and provides insights into how different actions relate to one another. Visualization of the structure of co-occurrences facilitates the meaning-making of behavioral data by facilitating the identification of action patterns. Finally, ENA allows us to quantitatively compare the action differences between different conditions.

Results

We applied ENA to model teachers' actions across the three conditions that we are interested in, i.e., no dashboard, mirroring and guiding conditions. Following a mixed-methods approach, we triangulated quantitative (log data) and qualitative data (post-activity questionnaire responses from teachers) to contextualize and produce results about the three conditions. Figure 6 shows the averaged networks generated for the six teachers' actions in the three different conditions. Figure 7 shows the distribution of teachers' actions in detail across the three conditions.

A visual inspection of the structures of the averaged networks presented in Fig. 6 shows there is a difference in teachers' actions in the no dashboard condition when compared to the mirroring and guiding conditions. The averaged networks generated for mirroring and guiding conditions have similar network structures (see Fig. 6b, c). However, the connection strengths (co-occurrences) between nodes are different.
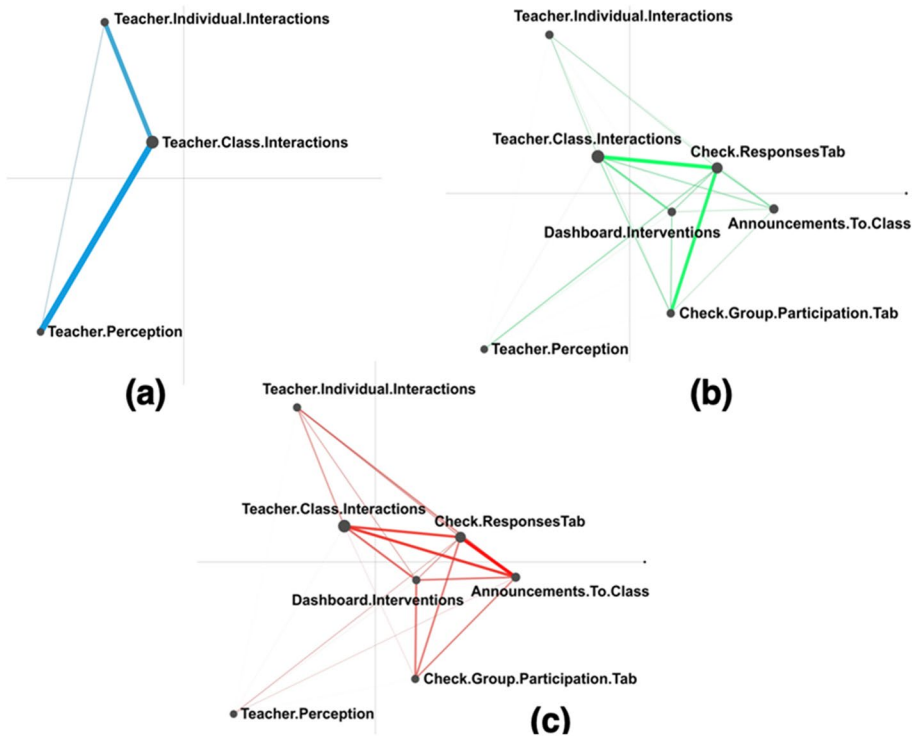
Fig.6 Mean networks of teachers' actions in the **a** no dashboard, **b** mirroring condition and **c** guiding condition representing the connections between different actions while orchestrating collaboration
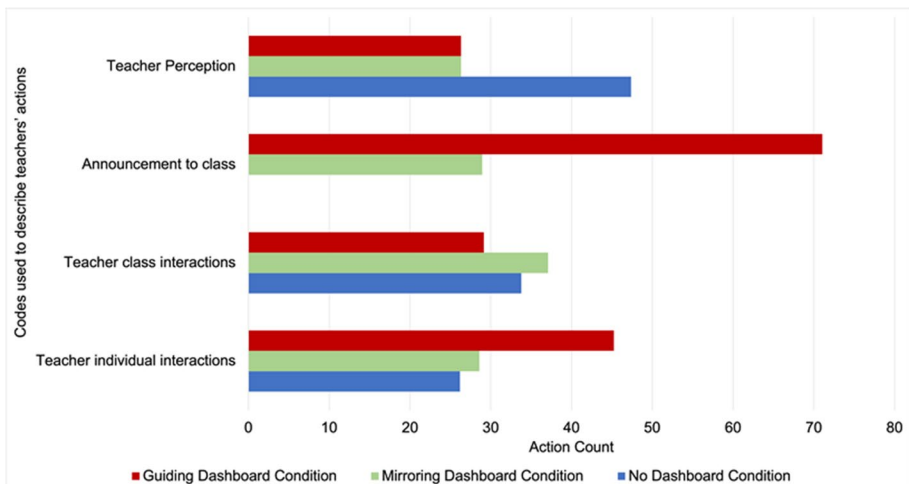

Fig. 7 A comparison of teacher's actions across the guiding, mirroring and no dashboard conditions

No dashboard condition

In the no dashboard condition strong co-occurrences between the following codes are visible: *teacher individual interactions* and *teacher class interactions*, *teacher class interactions* and *teacher perception* (see Fig. 6a). The observation of absent connections between the node that represents *announcements to class* and other nodes in the ENA diagram suggests that in this condition teachers were not in a position to make announcements to the class regarding time available, phase transitions of the script and student participation.

Figure 7 shows the distribution of teachers' actions across the three conditions (except dashboard related actions that are not common to all three conditions). As it can be seen in Fig. 7, a high number of *teacher perception* actions are visible in the no dashboard condition. Figure 8 shows the distribution of different types of *teacher class interactions* across the three conditions. As shown in Fig. 8, the most frequent *teacher class interactions* in the no dashboard condition were related to surveying activities. A high number of perception actions and *teacher class interactions* in the form of surveying indicate that in the no dashboard condition teachers put effort to diagnose the state of collaboration through perception and surveying.

The post-activity questionnaire responses collected from the teachers also confirmed the above results. Teachers mentioned that in the no dashboard condition it became impossible to follow the activity evolution over time, e.g., "I had to ask students several times if they had finished the activity". Teachers had less awareness over the state of the learning situation and faced problems in focusing on both epistemic and social aspects of the learning activity: e.g., "I was not aware whether students have problems in formulating answers. They all were silent. I couldn't make sure they were engaged in the task or they were doing something else". There were indications that they felt out of control: e.g., "Very difficult
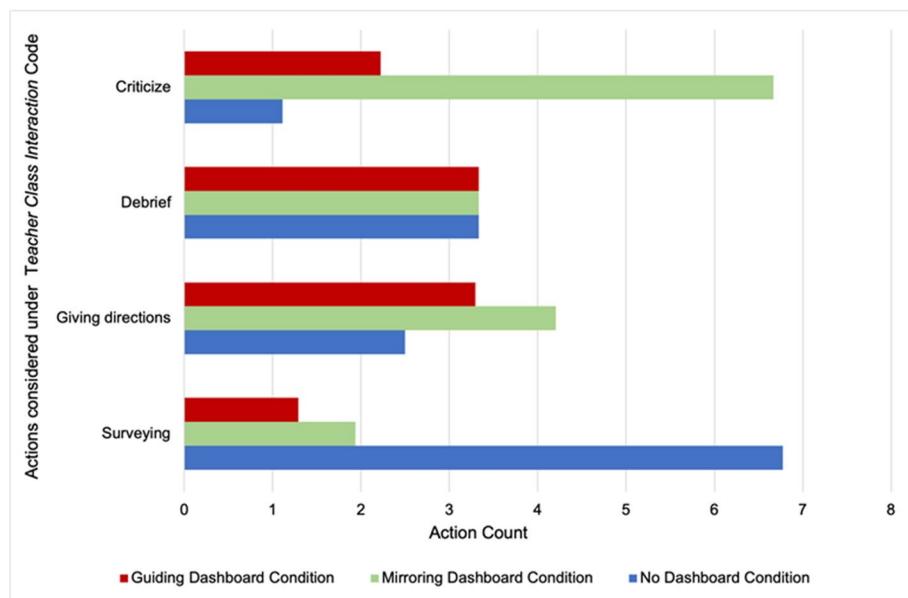

Fig. 8 An overview of teacher-class interaction actions across the guiding, mirroring and no dashboard conditions

to obtain the whole picture. I was stressed. I felt I did not have control over the activity".
Teachers' perception and classroom cues were used to gain awareness: e.g., "If everybody
is in silence this normally means they have started the activity"; "I know that the task is
done when the noise appears in the classroom…students start talking among them".

The above findings are in alignment with our expectations about the no dashboard condition (see "Introduction" section).

Mirroring condition

To disentangle the differences between the mirroring and guiding conditions, we generated
a difference network by subtracting the average connection strengths for teachers' actions
in the guiding condition from the average connection strengths for teacher actions in the
mirroring condition (see Fig. 9). Each line in Fig. 9 was colored to indicate which of the
two networks contains stronger co-occurrence.

Figure 9 denotes that the most frequent co-occurrences in the mirroring condition are
not the same as in the guiding condition. For instance, in the mirroring condition, stronger
edges exist between *teacher class interactions* and *check responses tab, teacher class interactions* and *check group participation tab, check responses tab* and *check group participation tab*. This implies that teachers frequently visited the information presented in the
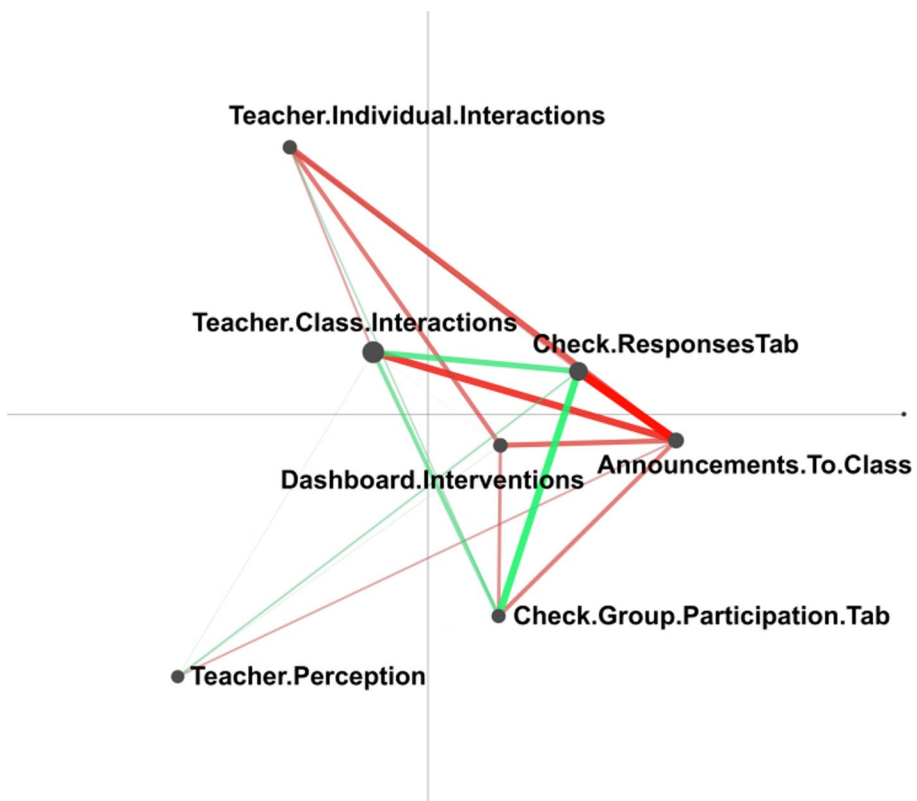

Fig. 9 Difference network for mirroring (in green) and guiding conditions (in red)

dashboard to evaluate the learning situations, which led them to take actions at the class level in the form of teacher-class interactions (see also Fig. 7). As shown in Fig. 8, a majority of *teacher class interactions* were in the form of criticizing lack of participation (6.6%) and providing further directions to students regarding the CSCL activity (4.2%). Moreover, the screen-captured data from the dashboard tablet showed that teachers consulted the information presented in the dashboard more often in the mirroring condition (137 times) when compared to the guiding condition (95 times). The *Responses Related* tab that provided information related to the epistemic aspect of the CSCL activity was consulted more often (the *Response Related tab* was selected 80 times and the *Participation Related tab* was selected 57 times).

Moreover, in the post-activity questionnaire teachers mentioned that having access to the dashboard helped to increase awareness and control collaboration "Design of the dashboard itself is user-friendly and intuitive. I had the opportunity to see all the answers. Overall picture of collaboration is provided". However, the teachers mentioned that in the mirroring condition they were mostly concentrated on one aspect of collaboration, namely, evaluating the content, and missed the chance to react to other aspects of the activity, e.g., changing activity duration, "On occasions I was concentrated on one aspect (e.g. reading their answers), I could not pay attention to other aspects in the dashboard (progress in the participation), so I missed elements to which I could have reacted, like adding more time in some phases". Finally, in the mirroring condition teachers made relatively fewer announcements (see Fig. 7) and conducted fewer dashboard interventions (see Table 4) when compared to the same actions in the guiding conditions.

The above results are in alignment with our expectations about the mirroring condition (see "Introduction" section) with the exception that teachers made fewer announcements to the class and focused more on the artefacts produced by the students.

Guiding condition

Figure 7 shows the presence of a high number of announcements in the guiding condition when compared to the mirroring condition. The stronger edges exist in Fig. 9 between *announcements to class* and *check response tab*, *announcements to class* and *check group participation tab*, *announcements to class* and *dashboard interventions, announcements to class* and *teacher perception*. These connections imply that in the guiding condition announcements were mostly informed by the information presented in

Table 4 Dashboard interventions

Intervention	Guiding condition	Mirroring condition
Posting messages to groups	14	4
Rating on behalf of low participating groups	1	1
Next level action	4	3
Increase time action	10 (9 based on alerts and 1 self-directed)	5
End action	2	3
Pause action	2 (1 based on alerts and 1 self-directed)	0

the *Responses Related* tab of the dashboard and somewhat informed by the information presented in the *Participation Related* tab of the dashboard, dashboard interventions and perceptions.

Further analysis conducted using log data suggests the presence of a high number of announcements in the guiding condition. For instance, during the answer submission phase of the pyramid script, five teachers received the "increase answer submission time" alert and one teacher received a "no keywords detected" alert. As a reaction to those alerts, teachers either increased the answer submission duration or paused the activity to give further guidelines. Teachers used this extra time to read students' answers and to check other statistics available, e.g., online-offline counts presented in the *Responses Related* tab of the dashboard (represented as a strong edge between *announcements to class* and *check response tab* codes in Fig. 9). The awareness gained by examining dashboard information led teachers to make announcements to the class about time remaining and activity participation. Table 5 provides an excerpt that exemplifies such connections.

Moreover, teachers also received alerts during the voting stages of the Pyramid script. For instance, increase time alerts were received by one teacher during the first voting level, and three teachers received the same alert during the second voting stage. All teachers reacted to these alerts, and as a result they made announcements to the class about remaining time and phase transitions. Teachers also referred to the information presented in the group *participation tab* of the dashboard to comment about activity participation. The strong edges that exist in Fig. 9 between *announcements to class* and the *check group participation tab* and between *announcements to class* and *dashboard interventions* in the guiding condition exemplify the aforementioned behaviors of the teachers. Furthermore, the connection between *announcements to class* and *teacher perception* reveals that some announcements were also influenced based on perception.

In the post-activity questionnaire, teachers mentioned that receiving alerts in the dashboard made necessary script changes (critical moments) upfront and put them in control, "I really felt I was in control. I could concentrate on those elements that interested me more (reading students' answers to identify misconceptions or issues of interest for later discussion). Even if I was not paying attention to activity participation and progression, the dashboard alerted me of critical moments in this respect", "The alerts shown by the system are very quick to read and do not disturb my tasks, they are helpful to react to certain moments of the activity".

However, teachers also mentioned that reacting to these alerts depended on the constraints of the classroom: e.g., "I decided to react to some of them, depending on other aspects of the context (like the overall time I could use for this activity). It is surprising that this happened to me even in a small group class. So, I guess this would be even more critical in larger classrooms". Moreover, in some situations teachers mentioned that receiving alerts about known information did not add value: e.g., "sometimes, I was carefully paying attention to dashboard information about activity progression, and I felt the alerts were a bit annoying – as offering information I already knew".

We also asked teachers opinions regarding the criteria used to generate alerts. Teachers highlighted some ideas that were not evaluated in the present study but proposed suggestions for future studies "I wonder if it is valid for activities where time expected for discussing and rating is long. In this case, half of the time allocated would not work but maybe ¾ of the time allocated, or this can be a parameter modifiable by the teacher". All six teachers agreed that alerts provided guidance to act and were useful to manage the activity. Teachers also mentioned they felt confident to react to alerts and the number of alerts shown in the dashboard was adequate (did not disturb orchestration).

Table 5 Excerpt from coded data in the guiding dashboard condition

Line	Teacher's observable behavior	Code
1	Teacher is scrolling and reading the answers submitted by the individual students to himself	Check.Response.Tab
2	"Think about the message given in the case study before submitting answers to the question"	Teacher.Class.Interaction
3	Lack of expected keywords in students' answers detected alert appeared on the dashboard and teacher paused the activity	Dashboard. Intervention
4	"Okay, the activity is paused"	Announcements.To.Class
5	Teacher provided more hints/instructions on how to write better answers to the given task	Teacher.Class.Interaction
6	"Okay, now I will resume the activity"	Announcements.To.Class
7	Teacher is scrolling and reading the answers submitted by the individual students to himself	Check.Response.Tab
8	Teacher again provided more hints/instructions on how to write better answers to the given task	Teacher.Class.Interaction
9	Teacher is scrolling and reading the answers submitted by the individual students to himself	Check.Response.Tab
10	"One minute left and we have only four answers so far"	Announcements.To.Class

Moreover, dashboard interventions are more prevalent in the guiding condition when compared to the mirroring condition (see Table 4). Some of these interventions were guided by the alerts, e.g., time and pause actions, whereas others were self-directed. Further, as shown in Fig. 7, in the guiding condition teachers engaged in fewer class interactions but more targeted interactions at the individual and at the group level (see Fig. 7) by answering questions from individual students and posting messages to groups. Log data showed that in the guiding condition teachers posted more messages to groups (14 times) when compared to the mirroring condition (4 times). The following predefined messages were posted to the groups, "Please rate the answers to finish the activity" (6 times), "I see that you're not discussing answers with your fellow group members" (7 times), "Have you already discussed your rating decisions with the fellow group members?" (1 time). We interpret the lack of criticism and surveying in the guiding condition as the result of teachers engaging in direct communication with problematic groups by posting messages. According to the post-activity questionnaire responses that teachers received, we observe alerts having stimulated their targeted interactions: "Alert made me aware that some groups require more time for voting and discussion. I also sent some messages to groups to motivate them to finish voting and when the delay of voting is due to misunderstandings about different answers submitted by students, I asked them to discuss". Although the overall teacher-class interactions have decreased in the guiding condition due to such targeted interventions (at the individual and group level), teachers did not reduced the essential classroom guidance in the form of directions for collaboration in the guiding condition (see Fig. 8).

The above findings are in alignment with our expectations about the guiding condition (see "Introduction" section).

Cognitive load

Finally, the differences between the three conditions were also evaluated based on the perceived cognitive load of the teachers (see "Methods" section). On average, in the guiding condition teachers reported a cognitive load of 6.2 (SD = 3.27). In the no dashboard condition, teachers reported a cognitive load of 5.6 (SD = 5.54), and the lowest value was reported for the mirroring condition, which was 5.4 (SD = 2.7).

Discussion

As expected, the six teachers participated in the study indicated that they had less awareness over the epistemic and social aspects of the learning activity in the no dashboard condition. ENA results and subjective responses of the teachers both confirmed that in this condition, they were out of control, and they could not make announcements to the class regarding time, phase transitions, and student participation during the activity. Given the small number of teachers participating in our study, we take these findings as early indications that having access to orchestration tools that visualize information about collaboration and provide a means for guiding collaboration become beneficial for orchestration (Wise & June 2019; Echeverria et al., 2018; Schwarz et al., 2021). The post-activity questionnaire responses from the teachers also confirmed that having access to the dashboard provided awareness and control regarding collaboration.

However, the results of the study indicated that the mirroring and guiding support provisions influenced teachers' orchestration actions differently. On the one hand, we expected that in both mirroring and guiding conditions teachers would focus equally on both epistemic and social aspects of the learning situation. However, in the mirroring condition they focused more on the epistemic aspect. The knowledge gained by making sense of the information presented in the dashboard (related to the epistemic aspect) led teachers to take action mostly in the form of teacher-class interactions (Verbat et al., 2014). On the other hand, we did not expect to see a difference in the number of announcements they would make under the mirroring and guiding support conditions. However, the results indicated that in the mirroring condition teachers made fewer announcements to the class (about important aspects related to time, script progression, and activity participation). Furthermore, in the mirroring condition teachers conducted several dashboard interventions even in the absence of alerts, indicating that they were able to interpret information and facilitate collaboration when required. However, when compared to the guiding condition, the number of dashboard interventions performed by the teachers remained low as expected. We also observed that they performed fewer targeted interventions in the mirroring condition when compared to the guiding condition.

The above findings could be related to the existence of a competitive workload that may add to the teachers' workload during orchestration as they engage in evaluating the content produced by students in real-time. This workload related to the epistemic aspect of collaboration, which could be referred to as content load. The content load is equally important in orchestrating collaboration, although previous studies have mostly concentrated on supporting teachers to solve group assignment issues and degree of student participation (Berland et al., 2015; Duque et al., 2015).

Moreover, in our study the CSCL activities proposed by teachers were open-ended in nature, i.e., there were no simple "yes/no" answers to a given problem, but instead multiple correct solutions and elaborations were possible. We think that evaluating artefacts produced by students to open-ended tasks that do not have simple "yes/no" answers contributes to the creation of a high content load for the teachers. The fact that teachers that consult the information presented in the *Responses Related* tab of the dashboard more often check students' artefacts confirms our interpretation. As a consequence of teachers spending more time evaluating students' artefacts, they ended up having less time to attend to problems at both the individual and group levels. This resulted in interactions at the class level as opposed to targeted interactions at individual and group levels. Moreover, as teachers focused more on the artefacts produced by the students, they may have fewer cognitive resources available to process other relevant aspects such as time available, degree of activity participation, and script progressions. Lack of focus towards such aspects has resulted in fewer announcements in the mirroring condition. This aligns with findings from previous research, which highlighted that teachers often miss critical aspects of collaboration as they pay more attention to assess students' artefacts (Martinez-Maldonado et al. 2015).

When compared to the mirroring condition, in the guiding condition, teachers conducted overall a higher number of dashboard interventions as expected. Some of these interventions were initiated as a reaction to the alerts, e.g., modification of time allocated to different script phases, yet some were self-directed, e.g., posting messages to groups. A possible explanation for a high number of self-directed interventions in the guiding condition is that the alerts informed critical moments and increased teachers' awareness regarding collaboration, which helped them to initiate other relevant interventions. Teachers also confirmed in the post-activity questionnaire that receiving alerts in the dashboard offloaded

constant monitoring of the activity and provided an opportunity to focus on interventions at individual and group levels.

As a result of teachers' reactions to alerts and self-directed actions, students were given more time to submit answers (during the answer submission phase of the script) and to evaluate answers from peers (during the voting phases of the script). The result was an overall fruitful collaborative learning situation. Teachers also used this additional time to evaluate the social and epistemic aspects of collaboration. For instance, by providing suggestions to improve answers (as a result of reacting to a *no keywords detected* alert) intervening in lower participation groups by sending messages or sometimes performing voting on behalf of the lower participation groups. Given the small number of teachers participating in our study, we take these findings as initial claims to indicate that guiding support is beneficial in orchestrating collaboration when compared to mirroring support.

These findings are in line with similar research conducted previously where it has been shown that alerts can provide additional support to increase teachers' awareness, saving them from constant monitoring (Tissenbaum & Slotta, 2019), prompting pedagogical actions that are conductive to learning (Schwarz et al., 2021), and providing additional help to control the flow of activity, i.e., script adaptations (Martinez-Maldonado et al. 2015). However, previous research has also shown that sometimes teachers disregard actions recommended using alerts as they have to prioritize other demands of the classroom such as limited time available to finish the activity (Martinez-Maldonado et al. 2015). Similar opinions were provided by the teachers who participated in our study as well.

The findings of the study that elucidated teachers' actionable differences under different support provisions shed light on how to deconstruct the notion of orchestration load into different facets: *situation evaluation, goal formation,* and *action-taking*. For instance, on the one hand, in the mirroring condition teachers attempted to evaluate the learning situation based on the information presented in the dashboard (*situation evaluation*). Although teachers may have had an overall picture of the learning situation, they were not explicitly supported to take actions as in the guiding condition *(action taking)*. Hence they had to constantly evaluate the learning situation and formulate goals *(goal formation).* As they employed their cognitive resources for *situation evaluation* and *goal formation* this might have reduced their ability to perform necessary dashboard interventions. This has resulted in an overall lower amount of dashboard interventions and orchestration actions in the mirroring condition when compared to the guiding condition. On the other hand, in the guiding condition alerts may have provided additional support for *situation evaluation, goal formation, and action-taking,* The additional support provided by the alerts for evaluating the learning situation and goal formation may have guided teachers to perform a high number of dashboard interventions when compared to the mirroring condition. However, in the guiding condition, when teachers were focusing on the content produced by the students, they were also informed of the need to take orchestration actions. Receiving alerts in the dashboard, while they were assessing the content in real-time, may have resulted in a situation in which teachers were required to distribute their attention to evaluate both epistemic and social aspects of collaboration. Focusing both on the epistemic aspect of collaboration as well as on the recommended actions simultaneously may have created a situation that is cognitively demanding. The competing nature of content load and orchestration load together with the high number of orchestration actions may have resulted in a slightly high cognitive load for the teachers in the guiding condition.

Nevertheless, it should be noted that cognitive load under different support conditions was measured on a scale from 1 to 20, and overall the recorded values for cognitive load are low in both mirroring and guiding conditions. Although we expected that teachers

might experience a high cognitive load in the mirroring condition and relatively low cognitive load in the guiding condition, the results show that in both conditions the cognitive load experienced by the teachers is low. This seemed to indicate that having access to the dashboard supported teachers who participated in our study to conduct orchestration actions and the information presented, and other design features of the dashboard did not overwhelm them cognitively.

Finally, regarding the presentation of the results in this study, it should be noted that ENA has provided a powerful technique to model differences in teacher behavior without losing the temporal order of the observed action sequences. ENA facilitated us to obtain a deeper understanding of how different actions co-occur as teachers' engage in regulating collaboration under different supporting conditions. As previously mentioned in the context of CSCL, many studies have focused on using ENA to model collaborative interactions of students yet to the best of our knowledge only a few studies have leveraged this technique to model teachers' orchestration actions.

Limitations of the study

One limitation of our study is the limited number of teachers who participated in it. Although conducting research studies with a limited number of teachers is common in teacher-oriented studies (Martinez-Maldonado, 2019; Wise & Jung, 2019), due to practical constraints, we acknowledge that the lower sample size reduces the generalizability of the results presented. Hence, the findings of the study need to be interpreted with caution, and we must treat it as an initial, exploratory study with a limited number of teachers. Although the sample size of our study is limited, we believe that the interpretation of their orchestration actions presented together with the ENA results is useful for setting expectations for potential future discoveries in bigger studies with larger samples of teachers.

Furthermore, the teachers who have participated in our experiments were all computer literate. All teachers had experience in using technology for their day-to-day teaching activities. However, it would be interesting to conduct further studies to explore how teachers with different backgrounds would use these types of tools in authentic settings. Recent studies have pointed out that teachers' data literacy, trust in technology, and experience may affect their use of LA tools (Verbert et al., 2020; Schwarz et al. 2021). Hence, conducting evaluation studies with teachers from different backgrounds and degrees of experience can enhance our understanding of whether the proposed orchestration technology is disruptive or reinforce positive change in their everyday practice.

Another issue is the counter-balancing of the three conditions in the within-subject design. Additional sequences would be required to achieve the full balance that a Latin Square design would provide, as an example. The design we adopted was a compromise due to the limited sample size, which nevertheless provided some means to reduce the bias related to having experienced a certain dashboard, and due to the number of teachers and sessions that were possible in the study. As all teachers who participated in our study had previous experience using PyramidApp, the possibility of having control of the pyramid script as such (based on their experience) was to a large extent common to all conditions.

Another limitation is that we have not used any eye-tracking software to track the exact information the teacher is looking at while using the dashboard. Although we have come up with codes such as *check responses tab* and *check group participation tab*, teachers may have also been looking at the time-related information or the dashboard controls presented at the top of the dashboard (not within a specific tab). We assumed that by switching tabs

the teacher is mainly observing the information presented within the particular tabs, not the information presented in the common space. However, incorporating eye-tracking software could have provided more precise details on teacher exploration of the dashboard information.

Moreover, we focused on collaborative learning activities that were scripted according to the Pyramid collaborative learning flow pattern. Even if we believe the pedagogical value of the Pyramid pattern and its application to multiple learning contexts, generalizing the findings of this study to other structures for learning activities requires further research.

Moreover, we have not reported in detail how teachers' actions affected students' activity engagement and learning gains. Although the primary focus of this study was related to supporting teachers in orchestrating classroom collaboration, we acknowledge that conducting pre-post test procedures to evaluate students' learning gains is important, and if done would have provided a more complete picture of the collaborative learning situation by closing the loop effectively (Clow, 2012). These aspects constitute limitations of our work and require further research.

Conclusions and future work

The notion of orchestration load is an important construct that needs to be given attention when designing tools and technologies to support orchestration. However, this notion is not sufficiently elaborated and differentiated in the existing literature as an important factor to be considered within the design decisions for teacher support tools. To this end, we investigated how different support provisions, i.e., mirroring and guiding, influence teachers' orchestration actions and how the presence and absence of certain orchestration actions under different support conditions can be explained by taking into account the notion of orchestration load. In that regard, we deconstructed the notion of orchestration load into three different facets, namely: situation evaluation, goal formation, and action-taking. We also identified other competing load aspects, i.e., content load, which emerges due to the real-time evaluation of the epistemic facet of the learning situations. The different facets of the orchestration load together with the competing content load can be used to interpret how different support provisions influence teachers' orchestration actions and to decide which types of support are beneficial for teachers in real-time. The findings of this study elucidated that guiding support assisted teachers in taking orchestration actions. The orchestration actions performed by the teachers were also found to benefit student collaboration, in contrast to the mirroring support, which led to fewer orchestration actions. The limitation encountered and the findings of the study have resulted in further research directions as listed below.

First, the type of task and time allocated for collaboration can impact teachers' orchestration actions. In our study, the tasks were mostly related to sharing knowledge. It would be interesting to conduct further studies to explore how different types of tasks proposed in different subject domains impact teachers' orchestration actions. Moreover, although teachers may value receiving alerts in the dashboard for activities planned for a shorter duration due to a high workload, this might differ for activities that are planned for longer durations. In such situations, teachers may have enough time to interpret observed patterns of collaboration and to take appropriate actions even without the support of alerts. Further research is needed to address those aspects going forward.

Second, we have noticed that there are subtle differences in the perceived cognitive load of the teachers under different support conditions. Further studies with a bigger sample of teachers can shed light on how cognitive load changes in relation to different support provisions.

Moreover, further studies around measuring the notion of orchestration load are required. As stated previously, we assume that the existing studies refer to the notion of orchestration load as a black box due to the difficulties associated with limited research instruments to quantify this construct. It would be useful to carry out further studies to develop instruments to estimate orchestration load in a more nuanced way, taking into account the different facets proposed in this study, i.e., *situation evaluation, goal formation, and action-taking.*

We conclude by outlining design recommendations for teacher-facing dashboards based on the findings of the study and the lessons learned. First, as we have elaborated in this study and have also been proposed in previous research (Martinez-Maldonado et al., 2015; Schwarz et al., 2021), we recommend generating alerts to inform teachers about critical moments related to both epistemic and social aspects of collaboration. The criteria for generating alerts can be decided following a human-centered design perspective where teachers' as the main stakeholders of these types of tools need to engage in the design process and are given a voice to express context-specific knowledge and expertise (Dimitriadis et al., 2021). Moreover, in our study teachers proposed the importance of customizing the criteria for generating alerts according to the unique needs of their sessions. We suggest that such preferences can be configured by teachers along with the learning design parameters, which can later be translated to rules for generating personalized alerts tailored to the unique needs of particular learning situations. Not only the alerts but also the information presented in the teacher-facing dashboards can be particularized to match with teachers' preferences, hence producing customized dashboards. In the future, we are planning to address the aforementioned research directions.

References

Amarasinghe, I., Hernández-Leo, D., Michos, K., & Vujovic, M. (2020). An actionable orchestration dashboard to enhance collaboration in the classroom. *IEEE Transactions on Learning Technologies, 13*(4), 662–675.

Berland, M., Davis, D., & Smith, C. P. (2015). AMOEBA: Designing for collaboration in computer science classrooms through live learning analytics. *International Journal of Computer-Supported Collaborative Learning, 10*, 425–447.

Clow, D. (2012). The learning analytics cycle: Closing the loop effectively. In *Proceedings of the 2nd international conference on learning analytics and knowledge (LAK 2012)* (pp. 134–138). https://doi.org/10.1145/2330601.2330636

Csanadi, A., Eagan, B., Kollar, I., Shaffer, D. W., & Fischer, F. (2018). When coding-and-counting is not enough: Using epistemic network analysis (ENA) to analyze verbal data in CSCL research. *International Journal of Computer-Supported Collaborative Learning, 13*(4), 419–438. https://doi.org/10.1007/s11412-018-9292-z

Cuendet, S., Bonnard, Q., Do-Lenh, S., & Dillenbourg, P. (2013). Designing augmented reality for the classroom. *Computers and Education, 68*, 557–569.

Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. In P. A. Kirschner (Ed.), Three worlds of CSCL. Can we support CSCL? (pp. 61–91). Heerlen, Open Universiteit Nederland. Accessed 07/06/2021 Retrieved from https://telearn.archives-ouvertes.fr/hal-00190230

Dillenbourg, P. (2015). *Orchestration graphs: Modelling scalable education* (1st ed.). EPFL Press.

Dillenbourg, P., Järvelä, S., & Fischer, F. (2009). The evolution of research on computer-supported collaborative learning. In E. Duval, M. Sharples, & R. Sutherland (Eds.), *Technology-enhanced learning* (pp. 3–19). Springer.

Dillenbourg, P., & Jermann, P. (2010). Technology for classroom orchestration. In *New science of learning* (pp. 525–552).

Dillenbourg, P., & Tchounikine, P. (2007). Flexibility in macro-scripts for computer-supported collaborative learning. *Journal of Computer Assisted Learning, 23*(1), 1–13. https://doi.org/10.1111/j.1365-2729.2007.00191.x

Dillenbourg, P., Zufferey, G., Alavi, H., Jermann, P., Do-Lenh, S., Bonnard, Q., Cuendet, S., & Kaplan, F. (2011). Classroom orchestration: The third circle of usability. In *Proceedings of the international conference on computer-supported collaborative learning (CSCL'11)* (pp. 510–517).

Dimitriadis, Y., Martínez-Maldonado, R., & Wiley, K. (2021). Human-centered design principles for actionable learning analytics. In *Research on E-learning and ICT in education: Technological, pedagogical and instructional perspectives* (pp. 277–296).

Duque, R., Gomez-Perez, D., Nieto-Reyes, A., & Bravo, C. (2015). Analyzing collaboration and interaction in learning environments to form learner groups. *Computers in Human Behavior, 47*, 42–49. https://doi.org/10.1016/j.chb.2014.07.012

Echeverria, V., Martinez-Maldonado, R., Shum, S. B., Chiluiza, K., Granda, R., & Conati, C. (2018). Exploratory versus explanatory visual learning analytics: Driving teachers' attention through educational data storytelling. *Journal of Learning Analytics, 5*(3), 72–97.

Gehlen-Baum, V., Weinberger, A., Pohl, A., & Bry, F. (2014). Technology use in lectures to enhance students' attention. In *Proceedings of the 9th European conference on technology enhanced learning* (pp. 125–137). Springer, Cham.

Herder, T., Swiecki, Z., Fougt, S. S., Tamborg, A. L., Allsopp, B. B., Shaffer, D. W., & Misfeldt, M. (2018). Supporting teachers' intervention in students' virtual collaboration using a network based model. In *Proceedings of the 8th international conference on learning analytics and knowledge* (pp. 21–25). https://doi.org/10.1145/3170358.3170394

Hernández-Leo, D., Asensio-Pérez, J. I., Dimitriadis, Y., & Villasclaras, E. D. (2019). Generating CSCL scripts: From a conceptual model of pattern languages to the design of real scripts. In P. Goodyear & S. Retalis (Eds.), *Technology-enhanced learning, design patterns and pattern languages* (pp. 49–64). Sense Publishers.

Knight, S., Wise, A. F., & Chen, B. (2017). Time for change: Why learning analytics needs temporal analysis. *Journal of Learning Analytics, 4*(3), 7–17.

Kobbe, L., Weinberger, A., Dillenbourg, P., Harrer, A., Hämäläinen, R., Häkkinen, P., & Fischer, F. (2007). Specifying computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning, 2*(2–3), 211–224. https://doi.org/10.1007/s11412-007-9014-4

Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing pedagogical action: Aligning learning analytics with learning design. *American Behavioral Scientist, 57*(10), 1439–1459.

Ma, L., Matsuzawa, Y., & Scardamalia, M. (2019). Using epistemic network analysis to ExploreWays of contributing to knowledge building discourse. In *Proceedings of the 13th international conference on computer-supported collaborative learning* (pp. 881–882). Lyon, France: International Society of the Learning Sciences.

Manathunga, K., & Hernández-Leo, D. (2018). Authoring and enactment of mobile pyramid-based collaborative learning activities. *British Journal of Educational Technology, 49*(2), 262–275. https://doi.org/10.1111/bjet.12588

Martinez-Maldonado, R. (2019). A handheld classroom dashboard: Teachers' perspectives on the use of real-time collaborative learning analytics. *International Journal of Computer-Supported Collaborative Learning, 14*(3), 383–411. https://doi.org/10.1007/s11412-019-09308-z

Martinez-Maldonado, R., Clayphan, A., Yacef, K., & Kay, J. (2015). MTFeedback: Providing notifications to enhance teacher awareness of small group work in the classroom. *IEEE Transactions on Learning Technologies, 8*(2), 187–200.

Oshima, J., Oshima, R., Ohsaki, A., & Splichal, J. (2019). Collective knowledge advancement through shared epistemic agency: Socio-semantic network analyses. In *Proceedings of the 13th international conference on computer supported collaborative learning* (pp. 57–64). Lyon, France: International Society of the Learning Sciences.

Prieto, L. P., Rodríguez-Triana, M. J., Martínez-Maldonado, R., Dimitriadis, Y., & Gašević, D. (2019). Orchestrating learning analytics (OrLA): Supporting inter-stakeholder communication about adoption of learning analytics at the classroom level. *Australasian Journal of Educational Technology, 35*(4), 14–33.

Prieto, L. P., Sharma, K., & Dillenbourg, P. (2015). Studying teacher orchestration load in technology enhanced classrooms. In G. Conole, T. Klobučar, C. Rensing, J. Konert, & É. Lavoué (Eds.), *Design for teaching and learning in a networked world* (pp. 268–281). Springer.

Prieto, L. P., Sharma, K., Kidzinski, Ł., & Dillenbourg, P. (2018). Orchestration load indicators and patterns: In-the-wild studies using mobile eye-tracking. *IEEE Transactions on Learning Technologies, 11*(2), 216–229. https://doi.org/10.1109/TLT.2017.2690687

Radkowitsch, A., Vogel, F., & Fischer, F. (2020). Good for learning, bad for motivation? A meta-analysis on the effects of computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning, 15*, 5–47. https://doi.org/10.1007/s11412-020-09316-4

Reimann, P. (2009). Time is precious: Variable-and event-centred approaches to process analysis in CSCL research. *International Journal of Computer-Supported Collaborative Learning, 4*(3), 239–257. https://doi.org/10.1007/s11412-009-9070-z

Rodríguez-Triana, M. J., Martínez-Monés, A., Asensio-Pérez, J. I., & Dimitriadis, Y. (2015). Scripting and monitoring meet each other: Aligning learning analytics and learning design to support teachers in orchestrating CSCL situations. *British Journal of Educational Technology, 46*(2), 330–343. https://doi.org/10.1111/bjet.12198

Roschelle, J., Dimitriadis, Y., & Hoppe, U. (2013). Classroom orchestration: Synthesis. *Computers and Education, 69*, 523–526. https://doi.org/10.1016/j.compedu.2013.04.010

Rummel, N., & Spada, H. (2005). Learning to collaborate: An instructional approach to promoting collaborative problem solving in computer-mediated settings. *Journal of the Learning Sciences, 14*(2), 201–241. https://doi.org/10.1207/s15327809jls14022

Saint, J., Gašević, D., Matcha, W., Uzir, N. A. A., & Pardo, A. (2020). Combining analytic methods to unlock sequential and temporal patterns of self-regulated learning. In *Proceedings of the 10th international conference on learning analytics and knowledge* (pp. 402–411). https://doi.org/10.1145/3375462.3375487

Schell, J., Lukoff, B., & Mazur, E. (2013). *Catalyzing learner engagement using cutting-edge classroom response systems in higher education: Classroom response systems and mediated discourse technologies*. Emerald Group Publishing Limited.

Schwarz, B. B., Swidan, O., Prusak, N., & Palatnik, A. (2021). Collaborative learning in mathematics classrooms: Can teachers understand progress of concurrent collaborating groups? *Computers and Education, 165*, 104151.

Schwendimann, B. A., Rodriguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., & Dillenbourg, P. (2016). Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies, 10*(1), 30–41. https://doi.org/10.1109/TLT.2016.2599522

Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *Journal of Learning Analytics, 3*(3), 9–45.

Sharples, M. (2013). Shared orchestration within and beyond the classroom. *Computers and Education, 69*, 504–506.

Shum, B., Ferguson, R., & Martinez-Maldonado, R. (2019a). Human-centered learning analytics. *Journal of Learning Analytics, 6*(2), 1–9.

Shum, S. B., Echeverria, V., & Martinez-Maldonado, R. (2019b). The multimodal matrix as a quantitative ethnography methodology. In *Proceedings of the international conference on quantitative ethnography* (pp. 26–40). https://doi.org/10.1007/978-3-030-33232-7_3

Siebert-Evenstone, A. L., Irgens, G. A., Collier, W., Swiecki, Z., Ruis, A. R., & Shaffer, D. W. (2017). In search of conversational grain size: Modelling semantic structure using moving stanza windows. *Journal of Learning Analytics, 4*(3), 123–139.

Soller, A., Martínez, A., Jermann, P., & Muehlenbrock, M. (2005). From mirroring to guiding: A review of state of the art technology for supporting collaborative learning. *International Journal of Artificial Intelligence in Education, 15*(4), 261–290.

Sung, H., Cao, S., Ruis, A., & Shaffer, D. W. (2019). Reading for breadth, reading for depth: Understanding the relationship between reading and complex thinking using epistemic network analysis. In *Proceedings of the 13th international conference on computer supported collaborative learning* (pp. 376–383). International Society of the Learning Sciences.

Sweller, J. (2020). Cognitive load theory and educational technology. *Educational Technology Research and Development, 68*(1), 1–16.

Tissenbaum, M., & Slotta, J. (2019). Supporting classroom orchestration with real-time feedback: A role for teacher dashboards and real-time agents. *International Journal of Computer Supported Collaborative Learning, 14*, 325–351.

van Leeuwen, A. (2015). Learning analytics to support teachers during synchronous CSCL: Balancing between overview and overload. *Journal of Learning Analytics, 2*(2), 138–162.

van Leeuwen, A., & Rummel, N. (2019). Orchestration tools to support the teacher during student collaboration: A review. *Unterrichtswissenschaft, 47*(2), 143–158.

van Leeuwen, A., & Rummel, N. (2020). Comparing teachers' use of mirroring and advising dashboards. In *Proceedings of the 10th international conference on learning analytics and knowledge (LAK'20)* (pp. 26–34). https://doi.org/10.1145/3375462.3375471

Verbert, K., Govaerts, S., Duval, E., Santos, J. L., Van Assche, F., Parra, G., & Klerkx, J. (2014). Learning dashboards: An overview and future research opportunities. *Personal and Ubiquitous Computing, 18*(6), 1499–1514.

Verbert, K., Ochoa, X., De Croon, R., Dourado, R. A., & De Laet, T. (2020). Learning analytics dashboards: the past, the present and the future. In *Proceedings of the 10th international conference on learning analytics and knowledge* (pp. 35–40). https://doi.org/10.1145/3375462.3375504

Wise, A. F., & Jung, Y. (2019). Teaching with analytics: Towards a situated model of instructional decision-making. *Journal of Learning Analytics, 6*(2), 53–69. https://doi.org/10.18608/jla.2019.62.4